Available at www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT®
ELSEVIER Stochastic Processes and their Applications 109 (2004) 189–201
stochastic processes and their applications
www.elsevier.com/locate/spa

Some inequalities for p-variations of martingales

Ming Fan*

Institute of Mathematics, Natural Sciences and Engineering, Dalarna University, Borlänge S-781 88, Sweden

Received 7 July 2003; received in revised form 25 September 2003; accepted 25 September 2003

Abstract

Some known inequalities concerning p-variations and conditional p-variations for discrete parameter martingales are sharpened and carried over in the more general context.

MSC: 60G42; 60G46; 46E30

Keywords: Conditional p-variation; Martingale inequality; p-variation; Rearrangement invariant; Sharp operator

1. Introduction and preliminaries

The main purpose of this paper is to study some inequalities concerning p-variations and conditional p-variations for discrete parameter martingales. Some known results, which were given in Burkholder (1973, 1991), Garsia (1973), Hitczenko (1990a), Wang (1991) and Weisz (1995), will be sharpened or carried over in the more general context.

Let (Ω, F, P) be a complete resonant probability space with the filtration $\{\mathsf{F}_n\}_{n\geqslant 0}$ for which $\mathsf{F} = \bigvee_{n\geqslant 0} \mathsf{F}_n$. The expectation operator and the conditional expectation operators relative to F_n are denoted by $E = E_0$ and E_n for $n \geqslant 1$, respectively. A sequence $(\omega_n)_{n\geqslant 0}$ of random variables is said to be adapted (resp. predictable) if ω_n is F_n- (resp. F_{n-1}-) measurable for all $n \geqslant 0$ (resp. $n \geqslant 1$). For each random variable, $f \in L^1 = L^1(\Omega, \mathsf{F}, P)$ with $Ef = 0$, we consider the corresponding martingale $f = (f_n)_{n\geqslant 0}$, where $f_n = E_n f$. The martingale differences are defined by

$$d_0 = d_0 f = 0 \quad \text{and} \quad d_n = d_n f = f_n - f_{n-1} \ (n \geqslant 1).$$

* Tel.: +46-23-778-853; fax: +46-23-778-054.
E-mail address: fmi@du.se (M. Fan).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2003.09.010

Sometimes we simply write $d=(d_n)_{n\geqslant 0}$. For stopping times λ and τ, the martingale f started at λ and stopped at τ is denoted by ${}^{\lambda}f^{\tau}$, and the martingale f stopped at τ is denoted by f^{τ}. In particular, $f^n=(f_0,f_1,\ldots,f_n,f_n,\ldots)$ is the corresponding martingale stopped at n.

Following Burkholder and Gundy (1970) and Weisz (1995), we consider some martingale operators T that map the set of martingales into the set of non-negative F-measurable functions. These kinds of operators satisfy the following conditions:

(B1) T is subadditive, i.e. if f,g and h are martingales for which $f=g+h$, then

$$T(f^n)\leqslant T(g^n)+T(h^n).$$

This property holds even for infinite sums.

(B2) T is homogeneous, i.e. $T(cf)=|c|T(f)$ for any constant c.

(B3) T is local, i.e. $T(f)=0$, if $\sum_{k=1}^{\infty}E_{k-1}\,|d_kf|^2=0$.

(B4) T is symmetric, i.e. $T(f)=T(-f)$.

We define $T_n(f)=T(f^n)$ for $n\geqslant 0$, $\tilde{T}(f)=\sup_{n\geqslant 0}T_n(f)$ and suppose that $T_0(f)=0$. An operator T is said to be adapted (resp. predictable) if $T_n(f)$ is F_n- (resp. F_{n-1}-) measurable for all martingales f and for all $n\geqslant 0$ (resp. $n\geqslant 1$). As examples, we consider the following special martingale operators: for a martingale $f=(f_n)_{n\geqslant 0}$, the maximal function $M(f)$, the p-variation $S_p(f)$ and the conditional p-variation $s_p(f)$, for $1\leqslant p<\infty$, are defined by

$$M(f)=\sup_{k\geqslant 0}|f_k|,\quad S_p(f)=\left(\sum_{k=1}^{\infty}|d_kf|^p\right)^{1/p},\qquad s_p(f)=\left(\sum_{k=1}^{\infty}E_{k-1}\,|d_kf|^p\right)^{1/p}.$$

Observe that all these operators satisfy the conditions (B1)–(B4); moreover, M and S_p are adapted, and s_p is predictable. Furthermore, we define the sharp operators $S_p^{\#}$ and $s_p^{\#}$ by

$$S_p^{\#}(f)=\sup_{n\geqslant 1}\left(E_n\left(\sum_{k=n}^{\infty}|d_k|^p\right)\right)^{1/p}$$

and

$$S_p^{\#}(f)=\sup_{n\geqslant 1}\left(E_n\left(\sum_{k=n+1}^{\infty}E_{k-1}(|d_k|^p)\right)\right)^{1/p}.$$

It is clear that $S_p^{\#}$ and $s_p^{\#}$ also satisfy conditions (B1)–(B4). In addition,

$$S_p^{\#}(f)\leqslant S_p(f)\quad\text{and}\quad s_p^{\#}(f)\leqslant s_p(f).$$

For an arbitrary martingale f, the following inequalities are well known:

$$\|S_p(f)\|_r\leqslant A_{p,r}\|s_p(f)\|_r\quad(0<r\leqslant p),\tag{1.1}$$

$$A_{p,r}\|s_p(f)\|_r\leqslant\|S_p(f)\|_r\quad(r\geqslant p),\tag{1.2}$$

$$\|S_p(f)\|_r \leqslant B_{p,r}\|S_p^{\#}(f)\|_r \quad (r \geqslant p), \tag{1.3}$$

$$\|s_p(f)\|_r \leqslant B_{p,r}\|s_p^{\#}(f)\|_r \quad (r \geqslant p). \tag{1.4}$$

Here $\|\cdot\|_r$ is the norm on $L^r(\Omega, \mathsf{F}, P)$, and $A_{p,r}$ and $B_{p,r}$ are constants only depending on p and r. We refer to Weisz (1995) for more details. Moreover, if $p = 2$, then the best constants in (1.1)–(1.3) are $A_{2,r} = \sqrt{2/r}$ (Wang, 1991) and $B_{2,r} = \sqrt{r/2}$ (Garsia, 1973). In Hitczenko (1990a), a martingale version of Rosenthal's inequality is given as follows:

$$\|M(f)\|_r \leqslant \frac{Cr}{\log r}(\|s_2(f)\|_r + \|M(\omega)\|_r) \quad (2 \leqslant r < \infty), \tag{1.5}$$

where $(\omega_n)_n$ is a predictable sequence of random variables which dominates $(|d_n|)_n$, $r/\log r$ is the best possible rate, and C is an absolute constant. In the present paper, we are going to determine the best possible constants/rates in these inequalities for all relevant p. In Section 2, we will sharpen inequalities (1.1)–(1.4). In Section 3, we will formulate Rosenthal's inequality (1.5) for operators s_p and L^r-spaces with $1 \leqslant p \leqslant 2$ and $p \leqslant r < \infty$, and even for more general rearrangement-invariant function spaces. The proof of our results will exploit some ideas of Burkholder (1973, 1991), Garsia (1973), Hitczenko (1990a, b), Johnson and Schechtman (1989), and Wang (1991). Finally, we include some related topics concerning interpolation of martingale Hardy spaces.

2. Sharp martingale inequalities on L^r-spaces

For the estimates in inequalities (1.1)–(1.4), let us begin with the following two lemmas.

Lemma 2.1. *For $x, y \geqslant 0$ and $d > 0$, we have*

$$(y+d)^{r/p}\left(\frac{x+d}{y+d} - \frac{p}{r}\right) \leqslant y^{r/p}\left(\frac{x}{y} - \frac{p}{r}\right) \quad (0 < r \leqslant p)$$

and

$$(y+d)^{r/p}\left(\frac{p}{r} - \frac{x+d}{y+d}\right) \leqslant y^{r/p}\left(\frac{p}{r} - \frac{x}{y}\right) \quad (r \geqslant p).$$

Proof. These inequalities were proved in Wang (1991) for $p = 2$. Observe that $r/p = (2r/p)/2$. Thus, the more general inequalities remain true for all $r > 0$. □

Lemma 2.2. *If $0 < \alpha \leqslant 1$, then $\alpha t - t^\alpha + (1-\alpha) \geqslant 0$ for all $t \geqslant 0$.*

Proof. Let $\varphi(t) = \alpha t - t^\alpha + (1-\alpha)$. Then $\varphi'(t) = \alpha - \alpha t^{\alpha-1}$. Hence, $\varphi'(t) = 0$ iff $t = 1$. Since $\varphi(0) = 1 - \alpha \geqslant 0$, $\varphi(1) = 0$ and $\lim_{t\to\infty}\varphi(t) = \infty$, it implies that $\varphi(t) \geqslant 0$ for all $t \geqslant 0$. □

Theorem 2.3. *Let f be a martingale. Then*

$$\|S_p(f)\|_r \leqslant (p/r)^{1/p}\|s_p(f)\|_r \quad (0 < r \leqslant p) \tag{2.1}$$

and

$$(p/r)^{1/p}\|s_p(f)\|_r \leqslant \|S_p(f)\|_r \quad (r \geqslant p). \tag{2.2}$$

Proof. It is enough to investigate the inequality in (2.1). For $0 < r \leqslant p$, let

$$W(x,y) = \left(\frac{p}{r}\right)^{r/p} y^{r/p-1}\left(x^p - \frac{p}{r}\,y\right) \qquad \text{for } x, y \geqslant 0.$$

By Lemma 2.1, we obtain that $W((x+d)^{1/p}, y+d) \leqslant W(x^{1/p}, y)$. Moreover, if we set $\alpha = r/p$ and $t = rx^p/py$, then by Lemma 2.2, we have

$$\frac{r}{p}\left(\frac{rx^p}{py} - 1\right) \geqslant \left(\frac{r}{p}\right)^{r/p}\left(\frac{x^r}{y^{r/p}}\right) - 1,$$

and hence $x^r - ((p/r)y)^{r/p} \leqslant W(x,y)$. This implies, for $\delta > 0$, that

$$\begin{aligned}
E\left(S_p^r(f^{n+1}) - \left(\frac{p}{r}\,(s_p^p(f^{n+1}) + \delta^p)\right)^{r/p}\right) &\leqslant EW(S_p(f^{n+1}), s_p^p(f^{n+1}) + \delta^p) \\
&\leqslant EW((S_p^p(f^n) + |d_{n+1}^p|)^{1/p}, s_p^p(f^n) + \delta^p + E_n(|d_{n+1}^p|)) \\
&= E(E_n W((S_p^p(f^n) + |d_{n+1}^p|)^{1/p}, s_p^p(f^n) + \delta^p + E_n(|d_{n+1}^p|))) \\
&= EW((S_p^p(f^n) + E_n(|d_{n+1}^p|))^{1/p}, s_p^p(f^n) + \delta^p + E_n(|d_{n+1}^p|)) \\
&\leqslant EW(S_p(f^n), s_p^p(f^n) + \delta^p) \leqslant EW(S_p(f^1), s_p^p(f^1) + \delta^p) \leqslant 0
\end{aligned}$$

by repeating the process n times. Therefore, $\|S_p(f)\|_r \leqslant (p/r)^{1/p}\|s_p(f)\|_r$ for $0 < r \leqslant p$, by letting $\delta \to 0$ and $n \to \infty$. □

Theorem 2.4. *If $r \geqslant p$, then*

$$\|S_p(f)\|_r \leqslant (r/p)^{1/p}\|S_p^{\#}(f)\|_r \tag{2.3}$$

and

$$\|s_p(f)\|_r \leqslant (r/p)^{1/p}\|s_p^{\#}(f)\|_r. \tag{2.4}$$

Proof. Set $\alpha = 1 - p/r$. Then we have

$$\frac{\alpha}{1-\alpha}\,t - \frac{1}{1-\alpha}\,t^{\alpha} + 1 \geqslant 0$$

for all $t \geqslant 0$ by Lemma 2.2, and hence $((r/p-1))t^r - (r/p)t^{r-p} + 1 \geqslant 0$ for all $t \geqslant 0$. For functions $\varphi, \psi \geqslant 0$ in L^r with $\|\psi\|_r \neq 0$, let $t = \|\varphi\|_r/\|\psi\|_r$. Then

$$\left(\frac{r}{p}-1\right)\|\varphi\|_r^r - \frac{r}{p}\|\varphi\|_r^{r-p}\|\psi\|_r^p + \|\psi\|_r^r \geqslant 0.$$

Observe, by Hölder's inequality, that

$$E(\varphi^{r-p}\psi^p) \leqslant E(\varphi^r)^{(r-p)/r}E(\psi^r)^{p/r} \leqslant \|\varphi\|_r^{r-p}\|\psi\|_r^p.$$

This implies that

$$E\left(\left(\frac{r}{p}-1\right)\varphi^r - \frac{r}{p}\varphi^{r-p}\psi^p + \psi^r\right) \geqslant \left(\frac{r}{p}-1\right)\|\varphi\|_r^r - \frac{r}{p}\|\varphi\|_r^{r-p}\|\psi\|_r^p + \|\psi\|_r^r$$
$$\geqslant 0.$$

Consequently,

$$E(\varphi^r - \psi^r) \leqslant \frac{r}{p}E(\varphi^{r-p}(\varphi^p - \psi^p)). \tag{2.5}$$

For the operator S_p, let $h_k = S_p^{r-p}(f^k) - S_p^{r-p}(f^{k-1})$. Then h_k is F_k-measurable. According to (2.5), we obtain

$$\begin{aligned}E(S_p^r(f^n)) &= \sum_{\upsilon=1}^{n} E(S_p^r(f^\upsilon) - S_p^r(f^{\upsilon-1})) \leqslant \frac{r}{p}\sum_{\upsilon=1}^{n} E(S_p^{r-p}(f^\upsilon)(S_p^p(f^\upsilon) - S_p^p(f^{\upsilon-1})))\\
&\leqslant \frac{r}{p}\sum_{\upsilon=1}^{n}\sum_{k=1}^{\upsilon} E(h_k(S_p^p(f^\upsilon) - S_p^p(f^{\upsilon-1})))\\
&= \frac{r}{p}\sum_{k=1}^{n}\sum_{\upsilon=k}^{n} E(h_k(S_p^p(f^\upsilon) - S_p^p(f^{\upsilon-1})))\\
&= \frac{r}{p}\sum_{k=1}^{n} E(h_k(S_p^p(f^n) - S_p^p(f^{k-1}))).\end{aligned}$$

Since

$$\begin{aligned}E(h_k(S_p^p(f^n) - S_p^p(f^{k-1}))) &= E(h_kE_k(S_p^p(f^n) - S_p^p(f^{k-1})))\\
&= E\left(h_kE_k\left(\sum_{\upsilon=k}^{n}|d_\upsilon^p|\right)\right) \leqslant E(h_kS_p^\#(f)^p),\end{aligned} \tag{2.6}$$

this, together with Hölder's inequality again, implies that

$$E(S_p^r(f^n)) \leqslant \frac{r}{p}\sum_{k=1}^{n} E(h_k S_p^{\#}(f)^p) = \frac{r}{p} E(S_p^{r-p}(f^n) S_p^{\#}(f)^p)$$
$$\leqslant \frac{r}{p} E(S_p^r(f^n))^{(r-p)/r} E(S_p^{\#}(f)^r)^{p/r}.$$

Therefore, $\|S_p(f)\|_r \leqslant (r/p)^{1/p}\|S_p^{\#}(f)\|_r$ by a direct calculation.

The inequality $\|s_p(f)\|_r \leqslant (r/p)^{1/p}\|s_p^{\#}(f)\|_r$ can be obtained in a similar way. The only difference is that the function $h_k = s_p^{r-p}(f_k) - s_p^{r-p}(f_{k-1})$ is F_{k-1}-measurable instead. Thus the formula in (2.6) reads

$$E(h_k(s_p^{r-p}(f^n) - s_p^{r-p}(f^{k-1}))) = E(h_k E_{k-1}(s_p^{r-p}(f^n) - s_p^{r-p}(f^{k-1})))$$
$$= E\left(h_k E_{k-1}\left(\sum_{\upsilon=k}^{n} E_{\upsilon-1}(|d_\upsilon^p|)\right)\right) \leqslant E(h_k s_p^{\#}(f)^p),$$

which completes the proof. □

For $p=2$, the estimates in (2.1)–(2.2), resp., (2.3), were obtained by Wang (1991, Theorem, 1), resp., Garsia (1973, Theorem II.1.2). Here we developed the methods used by them to the more general index p. For the sharpness of these inequalities, we follow an approach of Burkholder (1991) and Wang (1991) as shown below.

Theorem 2.5. *The constants in inequalities* (2.1)–(2.4) *are best possible.*

Proof. Let M be the set of all simple real conditionally symmetric martingales $f = (f_n)_{n\geqslant 0}$ with the filtration $\mathsf{F}_n = \sigma(f_0, f_1, \ldots, f_n)$ over the probability space $([0,1), \mathsf{B}, \mathrm{d}t)$, where B is the Borel σ-algebra on $[0,1)$. Let $\beta_{p,r}$ be the best constant in (2.1). If we choose $f = (f_n)_{n\geqslant 0}$ as a martingale with independent differences $(d_n)_{n\geqslant 0}$, then we obtain that $\alpha_{p,r} \leqslant 1 \leqslant \beta_{p,r}$.

Let $0 < r \leqslant p$, $\alpha = r/p$ and $\beta = \beta_{p,r}^r$. For $x, y \in \mathbf{R}$, we define

$$U_1(x,y) = \sup_{f\in\mathsf{M}} \{E|S_p^p(f) + x|^\alpha - \beta E|s_p^p(f) + y|^\alpha\}.$$

Then $U_1(\lambda^p x, \lambda^p y) = |\lambda|^r U_1(x,y)$, $U_1(x,y) \geqslant |x|^\alpha - \beta|y|^\alpha$, and $U_1(0,1) \leqslant 0$. For $0 < a < 1$ and for $f, g \in \mathsf{M}$, let $h = (h_n)_{n\geqslant 0} \in \mathsf{M}$ be the splice of f and g with weight a as in the proof of Burkholder (1991, Theorem 2.1). That is,

$$h_{n+1}(t) = \begin{cases} f_n(t/a) & \text{if } 0 \leqslant t < a, \\ g_n((t-a)/(1-a)) & \text{if } a \leqslant t < 1. \end{cases}$$

If $x_1, x_2, y \in \mathbf{R}$, and $x = ax_1 + (1-a)x_2$, then we have

$$a(E|S_p^p(f)+x_1|^\alpha - \beta E|s_p^p(f)+y|^\alpha) + (1-a)(E|S_p^p(g)+x_2|^\alpha - \beta E|s_p^p(g)+y|^\alpha)$$
$$\leqslant E|S_p^p(h)+x|^\alpha - \beta E|s_p^p(h)+y|^\alpha,$$

and hence

$$aU_1(x_1, y) + (1-a)U_1(x_2, y) \leqslant U_1(x, y).$$

Let $b > \beta_{p,r}$ and $0 < a < 1 \wedge (1/b^p)$. Then we have

$$\frac{a}{2} U_1\left(\frac{1}{a}, \frac{1}{ab^p}\right) + \frac{a}{2} U_1\left(-\frac{1}{a}, \frac{1}{ab^p}\right) + (1-a)U_1\left(0, \frac{1}{ab^p}\right)$$
$$\leqslant U_1\left(0, \frac{1}{ab^p}\right) \leqslant U_1\left(0, \frac{1}{ab^p} - 1\right),$$

and

$$U_1\left(\pm\frac{1}{a}, \frac{1}{ab^p}\right) \geqslant \frac{1}{a^\alpha} - \frac{\beta}{a^\alpha b^r} > 0.$$

This implies that $U_1(0, 1-ab^p) - (1-a)U_1(0,1) > 0$, and hence

$$((1-ab^p)^\alpha - (1-a))U_1(0,1) > 0.$$

Now we have $h(a) = (1-ab^p)^\alpha - (1-a) < 0$ and $h(0) = 0$. Thus $h'(0) = -\alpha b^p + 1 < 0$. Therefore, $b > (p/r)^{1/p}$ and hence $\beta_{p,r} \geqslant (p/r)^{1/p}$.

Let $r \geqslant p$ and $\alpha = r/p$. For $x, y \in \mathbf{R}$, we define

$$U_2(x, y) = \sup_{f \in \mathsf{M}} \{E|s_p^p(f)+x|^\alpha - \beta^{-1}E|S_p^p(f)+y|^\alpha\},$$

$$U_3(x, y) = \sup_{f \in \mathsf{M}} \{E|S_p^p(f)+x|^\alpha - \beta E|(S_p^\#)^p(f)+y|^\alpha\},$$

$$U_4(x, y) = \sup_{f \in \mathsf{M}} \{E|s_p^p(f)+x|^\alpha - \beta E|(s_p^\#)^p(f)+y|^\alpha\},$$

where $\beta^{1/r}$ is the best constant in (2.j) for U_j $(j = 2, 3, 4)$. The sharpness of inequalities (2.2)–(2.4) can be obtained similarly. □

3. On variants of Rosenthal's inequality

We begin with a result about the upper bounds of martingales given by Hitczenko (1990b).

Lemma 3.1. *Let ψ be a non-negative, strictly increasing function on $\mathbf{R}^+$, and let T be a martingale operator satisfying the conditions* (B1)–(B4). *Suppose that for all martingales f with the property that both $(|d_n|)_n$ and $(T({}^{n-1}f^n))_n$ are dominated by a* *predictable sequence of random variables* $(\omega_n)_n$ *the following inequality holds*:

$$P(|f_n| > \lambda(\|\tilde{T}f\|_\infty \vee \|M(\omega)\|_\infty)) \leqslant a\exp(-b\psi(\lambda)), \quad \lambda > 0,$$

for some constants a, b. *Then, there is an absolute constant* C_T *such that the inequality*

$$\|f\|_r \leqslant C_T\psi^{-1}(r)(\|\tilde{T}f\|_r + \|M(\omega)\|_r)$$

holds true for all martingales f *as above and for* $r \geqslant r_0$.

As an application of this result, we now extend Rosenthal's inequality (1.5) to the conditional variation operator s_p. The key point of our proof is a version of Prokorov's "arcsinh" inequality for martingales, which was used by Hitczenko (1990a) for s_2 on L^r with $2 \leqslant r < \infty$. For a sequence of random variables $(\omega_n)_n$, let us denote $M(\omega) = \sup_{k\geqslant 0}|\omega_k|$ as in the martingale case.

Theorem 3.2. *Let* $1 \leqslant p \leqslant 2$ *and* $p \leqslant r < \infty$. *Then there is a constant* C_p *such that the inequality*

$$\|M(f)\|_r \leqslant C_p \frac{r}{\log r}(\|s_p(f)\|_r + \|M(\omega)\|_r)$$

holds true for any martingale f *and for any predictable sequence of random variables* $(\omega_n)_n$ *which dominates* $(|d_n|)_n$.

Proof. Let us assume that $|d_k| \leqslant M$ a.s., and $\|s_p(f)\|_\infty = K < \infty$. Following the proof of Hitczenko (1990a, Propositions 3.1 and 3.2), we set, for $c > 0$,

$$g_n = \exp\left(c\sum_{k=0}^{n} d_k - \frac{c}{M}\sinh cM \cdot s_p^p(f^n)\right).$$

Then $(g_n)_{n\geqslant 0}$ is a supermartingale since

$$E_n(\exp(cd_{n+1}) - 1) \leqslant E_n(c|d_{n+1}|\sinh c|d_{n+1}|) = E_n\left(c^2|d_{n+1}|^p \frac{\sinh c|d_{n+1}|}{c|d_{n+1}|^{p-1}}\right)$$

$$\leqslant E_n(c|d_{n+1}|^p)\frac{\sinh cM}{M^{p-1}}.$$

The last inequality holds because the function $\sinh x/x^{p-1}$ increases on $\mathbf{R}^+$ for $1 \leqslant p \leqslant 2$. This implies that

$$P\left(\left|\sum_{k=1}^{n} d_k\right| \geqslant \lambda\right) \leqslant \exp\left(-c\lambda + \frac{cK^p}{M^{p-1}}\sinh cM\right) \quad \text{for all } \lambda > 0.$$

If we choose $c = (1/M)\operatorname{arc\,sinh} M^{p-1}\lambda/2K^p$, then we obtain

$$P\left(\left|\sum_{k=1}^{n} d_k\right| \geqslant \lambda\right) \leqslant \exp\left(-\frac{\lambda}{2M}\operatorname{arc\,sinh}\left(\frac{M^{p-1}\lambda}{2K^p}\right)\right) \quad \text{for all } \lambda > 0.$$

Therefore,

$$P(|f_n| \geqslant \lambda(\|s_p(f)\|_\infty \vee \|d\|_\infty)) \leqslant \exp\left(-\frac{\lambda}{2M}\operatorname{arc\,sinh}\lambda\right)$$

$$\leqslant 2\exp\left(-\frac{\lambda}{2}\log\left(1+\frac{\lambda}{2}\right)\right)$$

for all $\lambda > 0$. According to Lemma 3.1, we have

$$\|M(f)\|_r \leqslant \frac{C_p r}{\log r}(\|s_p(f)\|_r + \|M(\omega)\|_r). \qquad \square$$

Around 1990, the well-known Burkholder–Davis–Gundy inequality was placed in the more general framework of rearrangement-invariant function spaces (r.i. spaces in short) by several authors independently in terms of the Boyd indices and the standard stopping time argument (Antipa, 1990; Johnson and Schechtman, 1989; Novikov, 1991). It would be natural to expect that other martingale inequalities are also valid in this situation. One usually considers r.i. spaces X with $\underline{\alpha}_X > 0$. Here we denote by $\underline{\alpha}_X$ and $\bar{\alpha}_X$ the lower and higher Boyd indices for X. We refer to Bennett and Sharpley (1988) for further information of r.i. spaces. The next result of this section is to extend the martingale version of Rosenthal's inequality to r.i. spaces.

It is helpful to begin with a version of Johnson and Schechtman (1989, Lemma 4 and Corollary 2).

Lemma 3.3. *Let X be an r.i. space over (Ω, F, P) with $\alpha = \underline{\alpha}_X > 0$, and let f and g be non-negative random variables in X. Suppose that $\beta, \delta, \varepsilon > 0$ with $\beta\varepsilon^\alpha < 1$ satisfy*

$$P(f > \beta\lambda, g \leqslant \delta\lambda) \leqslant \varepsilon P(f > \lambda) \quad \textit{for all } \lambda > 0,$$

then

$$\|f\|_X \leqslant \frac{\beta}{\delta(1-\beta\varepsilon^\alpha)}\|g\|_X.$$

Lemma 3.4. *If X is an r.i. space with $\underline{\alpha}_X > 0$, then there exists a constant c_X such that*

$$\left\|\sum_{k=0}^{\infty} E_{k-1}\omega_k\right\|_X \leqslant c_X \left\|\sum_{k=0}^{\infty}\omega_k\right\|_X$$

for all sequences of non-negative measurable functions $(\omega_k)_{k\geqslant 0}$ in X.

Theorem 3.5. *Let X be an r.i. space over (Ω, F, P) with $\underline{\alpha}_X > 0$ and let $1 \leqslant p < \infty$. Then the inequalities*

$$\|S_p(f)\|_X \leqslant c_{p,X}(\|s_p(f)\|_X + \|M(d)\|_X) \quad (1 \leqslant p < \infty), \tag{3.1}$$

$$\|M(f)\|_X \leqslant c_{p,X}(\|s_p(f)\|_X + \|M(d)\|_X) \quad (1 \leqslant p \leqslant 2) \tag{3.2}$$

hold true for all martingales f.

Proof. Without loss of generality, let us only consider the inequality (3.1). We will first show that, if f is a martingale with the property that $(|d_n|)_n$ is dominated by a predictable sequence of random variables $(\omega_n)_n$, then the following good λ inequality:

$$P(S_p(f) > \beta\lambda, s_p(f) \vee M(\omega) \leqslant \delta\lambda) \leqslant \frac{2\delta^p}{(\beta-\delta-1)^p} P(S_p(f) > \lambda)$$

remains true for all $\lambda > 0$, $\beta > 1$ and $0 < \delta < \beta - 1$. To be more precise, let us define three stopping times as below:

$$\mu = \inf\{n \mid S_p(f^n) > \lambda\},$$

$$v = \inf\{n \mid S_p(f^n) > \beta\lambda\},$$

$$\sigma = \inf\{n \mid s_p(f^n) > \delta\lambda \text{ or } \omega_{n+1} > \delta\lambda\}.$$

Let $h = (h_n)_{n\geqslant 0}$ be the martingale f started at μ and stopped at $v \wedge \sigma$. That is,

$$h_n = \sum_{k=1}^{n} I(\mu < k \leqslant v \wedge \sigma) d_k.$$

Observe that, on $\{\mu = \infty\} = \{S_p(f) \leqslant \lambda\}$, $s_p(h) = 0$; on $\{0 < \sigma < \infty\}$,

$$\begin{aligned} s_p(h)^p &\leqslant s_p(f^\sigma)^p P(S_p(f) > \lambda) \leqslant (s_p(f^{\sigma-1})^p + E_{\sigma-1}(d_\sigma^p)) P(S_p(f) > \lambda) \\ &\leqslant (\delta^p\lambda^p + \omega_\sigma^p) P(S_p(f) > \lambda) \leqslant 2\delta^p\lambda^p P(S_p(f) > \lambda) \end{aligned} \tag{3.3}$$

and on $\{v < \infty, \sigma = \infty\}$,

$$S_p(h) = S_p(f^v) - S_p(f^\mu) - |d_\mu| \geqslant (\beta - 1 - \delta)\lambda. \tag{3.4}$$

Observe that

$$\|S_p(h)\|_p^p = \|s_p(h)\|_p^p \leqslant 2\delta^p\lambda^p P(S_p(f) > \lambda)$$

by Weisz (1995, (25)) and (3.3). This, together with (3.4), implies that

$$\begin{aligned} &P(S_p(f) > \beta\lambda, s_p(f) \vee M(\omega) \leqslant \delta\lambda) \\ &\quad = P(v < \infty, \sigma = \infty) \leqslant P(S_p(h) \geqslant (\beta - \delta - 1)\lambda) \\ &\quad \leqslant \frac{1}{(\beta-\delta-1)^p\lambda^p} \|S_p(h)\|_p^p \leqslant \frac{2\delta^p}{(\beta-\delta-1)^p} P(S_p(f) > \lambda). \end{aligned}$$

Let $\alpha = \underline{\alpha}_X$ and $\varepsilon = 2\delta^p/(\beta - 1 - \delta)^p$. Choose $\beta > 1$ and $0 < \delta < \beta - 1$ such that

$$\frac{\delta\beta^{1/\alpha p}}{\beta - 1 - \delta} < \frac{1}{2^{1/p}}.$$

Then $\beta\varepsilon^\alpha < 1$ and hence

$$\|S_p(f)\|_X \leqslant \frac{\beta}{\delta(1 - \beta\varepsilon^\alpha)} (\|s_p(f)\|_X + \|M(\omega)\|_X)$$

by Lemma 3.3. The estimate $\|S_p(f)\|_X \leqslant c_{p,X}(\|s_p(f)\|_X + \|M(d)\|_X)$ follows from this inequality, Davis' decomposition of a martingale as in Burkholder (1973, Section 14), and Lemma 3.4. The proof is a straightforward adaption of that of Burkholder (1973, Theorem 15.1). □

In case $X = L^r$ $(1 \leqslant r < \infty)$, the inequality (3.1) was proved by Rosenthal (1970) and Burkholder (1973) for $p = 2$, and by Weisz (1995) for $1 < p < \infty$.

Remark 3.6. Similarly, we can generalize the strong p-variation of martingales to r.i. spaces. For $1 \leqslant p < \infty$, the strong p-variation of a sequence $x = (x_n)_{n\geqslant 0}$, denoted by $W_p(x)$, is defined as follows:

$$W_p(x) = \sup\{\|(x_{n_k} - x_{n_{k-1}})_k\|_p \mid 0 \leqslant n_0 \leqslant n_1 \leqslant \cdots\}.$$

In fact, if X is an r.i. space over (Ω, F, P) with $\underline{\alpha}_X > 0$, then the inequalities

$$\|W_p(f)\|_X \leqslant c_{p,X}\|S_p(f)\|_X \quad (1 \leqslant p < 2),$$

$$\|W_p(f)\|_X \leqslant c_{p,X}\|M(f)\|_X \quad (2 < p < \infty),$$

hold true for all martingales f in X. We leave the proof to the reader.

4. Remarks on interpolation of martingale Hardy spaces

In Weisz (1995), Weisz formulated several interpolation results on Hardy and BMO spaces generated by some martingale operators in terms of the classical real interpolation methods, and used these results to prove some martingale inequalities. In fact, the results of Weisz (1995) continue to hold true in the general setting of real interpolation methods in the sense of Brudnyi and Krugljak with a quasi-power parameter (Brudnyi and Krugljak, 1991). In the final section, we give a quick review of this approach without proof. We refer to Bergh and Löfström (1976) and Brudnyi and Krugljak (1991) for the background knowledge on interpolation theory. The equality between Banach spaces means isomorphic equivalence.

Let us assume that $\bar{X} = (X_0, X_1)$ is a Banach couple with $\Delta\bar{X} = X_0 \cap X_1$ and $\Sigma\bar{X} = X_0 + X_1$. For $t > 0$, the J- and K-functionals are given by

$$J(t,x;\bar{X}) = \|x\|_0 \vee (t\|x\|_1) \qquad \text{for } x \in \Delta\bar{X},$$
$$K(t,x;\bar{X}) = \inf\{\|x_0\|_0 + t\|x_1\|_1 \mid x = x_0 + x_1, x_j \in X_j\} \qquad \text{for } x \in \Sigma\bar{X}.$$

We now introduce Brudnyi–Krugljak's K- and J-methods as follows: let Φ be a Banach function space on $\mathbf{R}^+ = (0,\infty)$ such that

$$1 \wedge t \in \Phi \quad \text{and} \quad \int_0^\infty 1 \wedge (1/t)|f(t)|\,\frac{\mathrm{d}t}{t} < \infty \quad \text{for all } f \in \Phi,$$

then we define

$$K_\Phi(\bar{X}) = \{x \in \Sigma\bar{X} \mid \|x\|_{K_\Phi} = \|K(t,x;\bar{X})\|_\Phi < \infty\}$$

(Brudnyi and Krugljak, 1991, (3.3.1))

and define $J_\Phi(\bar{X})$ as the space of all those $x \in \Sigma\bar{X}$ which permits a canonical representation $x = \int_0^\infty u(t)\,\mathrm{d}t/t$ for a strongly measurable function $u : \mathbf{R}^+ \to \Delta\bar{X}$ with the norm

$$\|x\|_{J_\Phi} = \inf_u \|J(t,u(t);\bar{X})\|_\Phi < \infty \quad \text{(Brudnyi and Krugljak, 1991, (3.4.3))}$$

If $T : X_j \to Y_j$ $(j = 0,1)$ are quasi-linear and bounded, then

$$T : K_\Phi(\bar{X}) \to K_\Phi(\bar{Y}) \quad \text{and} \quad T : J_\Phi(\bar{X}) \to J_\Phi(\bar{Y})$$

are also bounded with the interpolation norm $\leqslant \|T\|_0 \vee \|T\|_1$. Now let us define the Calderón operator S by

$$(Sf)(t) = \int_0^\infty 1 \wedge (t/s) f(s)\,\frac{\mathrm{d}s}{s}.$$

The function space Φ is called a quasi-power parameter if S is bounded on Φ. In this case, we have the equivalence $J_\Phi(\bar{X}) = K_\Phi(\bar{X})$ for all Banach couples $\bar{X}$ (Brudnyi and Krugljak, 1991, Corollary 3.5.35).

Let X be an r.i. space over (Ω, F, P) with nontrivial Boyd indices $\underline{\alpha}_X, \bar{\alpha}_X$ satisfying $0 < 1/r_1 < \underline{\alpha}_X \leqslant \bar{\alpha}_X < 1/r_0 < 1$. By an argument in Fan (2002, Remark 4.4), one can construct a quasi-power function space Φ, for which

$$X = J_\Phi(L^{r_0}, L^{r_1}) = K_\Phi(L^{r_0}, L^{r_1}). \tag{4.1}$$

Furthermore, if T is a martingale operator satisfying (B1)–(B4), then we can define the martingale Hardy space H_X^T, which is generated by X and T, and consists of all martingales f for which $\|f\|_{H_X^T} = \|\tilde{T}(f)\|_X < \infty$. In particular, we write $H_p^T = H_{L^p}^T$ for $1 \leqslant p \leqslant \infty$. Now we can formulate the following interpolation result for martingale Hardy spaces: if $T = M, S_p$ or if T is predictable, then

$$H_X^T = J_\Phi(H_{r_0}^T, H_{r_1}^T) = K_\Phi(H_{r_0}^T, H_{r_1}^T) \tag{4.2}$$

by using (4.1) and the inequality

$$K(t,f;H_1^T,H_\infty^T) \leqslant c\int_0^t (\tilde{T}f)^*(s)\,\mathrm{d}s = cK(t,\tilde{T}f;L^1,L^\infty)$$

for any martingale f and for some constant c in terms of Milman (1981, Lemma 3.1) and Weisz (1995, Lemmas 2 and 3).

As a consequence of (4.2), many martingale inequalities, including (1.1)–(1.4), can be easily carried over from L^r-spaces to the more general r.i. spaces. The drawback of this approach, however, is that on the one hand the constants appearing in those inequalities are far from the best possible, and on the other, it is difficult to deal with the “end points” $r = \underline{\alpha}_X$ or $r = \bar{\alpha}_X$.

References

Antipa, A., 1990. Doob's inequality for rearrangement-invariant function spaces. Rev. Roumuine Math. Pures Appl. 35, 101–108.

Bennett, C., Sharpley, R., 1988. Interpolation of operators. In: Pure and Applied Mathematics, Vol. 129. Academic Press, New York.

Bergh, J., Löfström, J., 1976. Interpolation spaces. In: Grundlehren der Mathematischen Wisseaschatter, Vol. 223. Springer, Berlin/Heidelberg/New York.

Brudnyi, Y.A., Krugljak, N.Y., 1991. Interpolation Functors and Interpolation Spaces. North-Holland, Amsterdam.

Burkholder, D.L., 1973. Distribution function inequalities for martingales. Ann. Probab. 1, 19–42.

Burkholder, D.L., 1991. Explorations in martingale theory and its applications. Lecture Notes in Mathematics, Vol. 1464, Springer, Berlin, pp. 1–66.

Burkholder, D.L., Gundy, G.F., 1970. Extrapolation and interpolation of quasi-linear operators on martingales. Acta Math. 124, 250–304.

Fan, M., 2002. Commutators in real interpolation with quasi-power parameters. Abstracts Appl. Anal. 7, 239–258.

Garsia, M., 1973. Martingale inequalities. Seminar Notes on Recent Progress, Mathematics Lecture Notes Series, Benjamin, New York.

Hitczenko, P., 1990a. Best constants in martingale version of Rosenthal's inequality. Ann. Probab. 18, 1656–1668.

Hitczenko, P., 1990b. Upper bounds for the L_p-norms of martingales. Probab. Theory Related Fields 86, 225–238.

Johnson, W.B., Schechtman, G., 1989. Martingale inequalities in rearrangement invariant function spaces. Israel J. Math. 64, 267–275.

Milman, M., 1981. On interpolation of martingale L^p spaces. Indiana Univ. Math. J. 30, 313–318.

Novikov, I.Ya., 1991. Martingale inequalities in rearrangement invariant function spaces. In: Musielak, J. et al. (Eds.), Function spaces, Teubner-Texte zur Mathematik, Vol. 120. Teubner, Stuttgart, pp. 120–127.

Rosenthal, H.P., 1970. On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. Israel J. Math. 8, 273–303.

Wang, G., 1991. Sharp inequalities for the conditional square function of a martingale. Ann. Probab. 19, 1679–1688.

Weisz, F., 1995. Martingale operators and Hardy spaces generated by them. Studia Math. 114, 39–70.